

Available online at www.sciencedirect.com**ScienceDirect**

Procedia Computer Science 38 (2014) 124 – 127

Procedia
Computer Science

10th Italian Research Conference on Digital Libraries, IRCDL 2014

<e-codices. Virtual Manuscript Library of Switzerland>: a case analysis of good practices for building a virtual library and reunifying dispersed handwritten collections

Fabio Cusimano*

Officina di Studi Medievali, Via del Parlamento 32, Palermo 90133, Italy

Abstract

This essay will focus on a case study of good practices in the field of projection and circulation of digital contents for humanistic research, starting from handwritten sources: the <e-codices - Virtual Manuscript Library of Switzerland> project. As is stated in the <e-codices> website, “the goal of the <e-codices> project is to provide access to all medieval and selected early modern manuscripts of Switzerland via a virtual library”, virtually reunifying dispersed handwritten collections. This goal was made possible thanks to the TEI-P5 standard, which “enables the exchange of metadata with other digitization projects, meta-search engines, and portal sites”.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Peer-review under responsibility of the Scientific Committee of IRCDL 2014

Keywords: digitization; metadata; encoding; XML; TEI-P5

1. Main features of the project <e-codices – Virtual manuscript library of Switzerland>

<e-codices – Virtual manuscript library of Switzerland> is a project developed by the Institute for Medieval Studies of the University of Freiburg and directed by Professor Christoph Flüeler. It is part of the digital collections coordinated by <e-lib.ch>¹, the “Swiss electronic library”, which also includes <e-manuscripta.ch>, a portal devoted to handwritten sources digitized by Swiss libraries and archives, and <e-rara.ch>, a portal devoted to digital reproductions of Swiss publications spanning from the 15th to the 19th century.

* Corresponding author. Tel.: +39-(0)91-586314; fax: +39-(0)91-333121.

E-mail address: f.cusimano@officinastudimedievali.it

The <e-codices> project has many strong points: from the graphical point of view, the website presents a well designed GUI (*graphic user interface*) and simple graphics that are easy to use; it is able to provide users (manuscript scholars, librarians and general public) with updates and information on the objectives and progress of the project itself and sub-projects (thanks to a newsletter and an RSS feed), as well as providing access to its digitized manuscripts; furthermore, all contents are available in four languages (English, French, German and Italian).

In terms of the possibility of searching for data and information pertaining to the digitized samples, the <e-codices> website allows free access (it does not require registration on the portal) to 1145 digitized manuscripts from 46 different libraries (data updated to September 10, 2014): use of files and metadata is permitted in compliance with the Creative Commons License, which permits the use of individual images for non-commercial purposes in the areas of education and research.

Moreover, although the project's main focus is the creation of a Swiss virtual library collecting and allowing access to all medieval and selected early modern manuscripts in Switzerland, the digital archive also includes three Swiss manuscripts stored in Austria and Germany, as well as three parts of two other Swiss manuscripts which are conserved in France, Russia and in the USA; <e-codices> acts as an "aggregator" of digital specimens, performing their digitization directly so as to grant access to these sources beyond the libraries which house the precious originals. We can find a confirmation of the special features of <e-codices> as a "virtual library" in what Riccardo Ridi states: "la biblioteca virtuale è qualcosa di connesso in qualche modo alle risorse informative digitali, ma difficilmente identificabile in modo più preciso, tanto da essere compatibile con un ampio ventaglio di significati. La biblioteca digitale costituisce al tempo stesso un caso particolare di biblioteca virtuale (tanto che talvolta i due termini vengono impropriamente utilizzati come sinonimi) e il caso limite di biblioteca ibrida in cui documenti analogici e procedure e servizi manuali siano ridotti a zero."⁶

In terms of the opportunity² (which, sadly, is often merely nominal) both for a digital library and for a virtual library to represent a space in which the scholarly community may not only conduct individual research, but also share information and ideas, the <e-codices> portal has also created a tool which Latin Paleography and Codicology scholars, as well librarians as anyone interested in the subject, will certainly appreciate and find highly useful: the *Annotation tool*, developed upon a request by researchers and libraries. This tool (which requires prior log-in through the individual manuscript summary page) enables scholars to collaboratively add to the information available for each manuscript, giving rise to a collaborative process enhancing the bibliographical references and annotations pertaining to an entire manuscript, its binding, a single page or range of pages. <e-codices> encourages collaboration with researchers in the field of manuscript scholarship by requesting that users suggest their *desiderata*, manuscripts that are important to their research for possible digitization and inclusion on the <e-codices> website: in this way it could be possible to write a new scholarly description, or edit or expand an existing description, or could also involve a digitization-based research project that demonstrates the advantages of making manuscript resources digitally available online.

Thanks to the <e-codices> virtual library, manuscripts from historic collections (as, for example, the prestigious collection of the historical library of St. Gallen, with over 500 digitized manuscripts) are virtually reunified³ for access via Internet and researchers are given a new tool for studying manuscripts: the <e-codices> virtual library³ provides unified record lists for dispersed Swiss collections and immediate Internet access to high-quality digital reproductions, linking scholarly descriptions to all works available in the virtual library, to supply a full range of searchable bibliographic information.

The reference to "good practices" in the title of this paper, therefore, derives from a number of factors which, combined, have contributed to the creation of an effective project: from the high quality of the digital copies³ to their easy access through a functional online visualization tool; from the graphic interface design to the portal's metadata "population" which, along with the manuscript summaries contributed by users, ensure accessibility, traceability and interoperability^{4,5,6} of the data.

1.1. The crucial function of descriptive metadata

The most specific feature of <e-codices> is its focus on the assignment of metadata^{7,8,9} to digital contents, interoperability, and compliance with the XML TEI-P5 encoding system. In terms of the particular features and

technical and functional properties which should characterize a digital library, the added value of the <e-codices> project lies in its ability to put content interoperability into practice, and to guarantee access to information through a number of channels (also from mobile devices), also enabling the performance of targeted searches aimed at specific descriptive categories.

In this context a key-role is played by the value of XML and, simultaneously, of the features of TEI-P5. A particularly important segment of TEI-P5 *Guidelines* is Chapter 10, focusing on *Manuscript description*, to which the authors of <e-codices> have referred in encoding^{10,11,12,13,14,15} the descriptions of manuscripts. This specific module has been developed for the description of medieval handwritten sources in compliance with European cataloguing procedures; however, based on the metalanguage XML, it is easily adaptable for other specific uses (its very adaptability is one of the main reasons of the success of XML); the encoding of information linked to handwritten primary sources cannot prescind from the original bibliographical sources. Finally, another objective of this method aims to reclaim information contained in existing catalogues.

1.2. The TEI-P5 tags used in <e-codices>

Metadata include information on shelf marks, titles, dates, short characterizations and detailed scholarly descriptions, to build a full manuscript package for presentation on the website.

According to Chapter 10 of the TEI-P5 *Guidelines*, focused on *Manuscript description*, the entire tag set issued by the TEI Consortium was adopted by the <e-codices> staff for the development of manuscripts' descriptions.

All these tags can make the retrieval of different types of data relating to manuscripts possible, thanks to the specialization of the tag set: the scholarly manuscript descriptions can be searched either using "Full Text" or using specific access points (*Author*, *Manuscript title*, *Text title*, *Place of origin*, *Incipit*, *Explicit* and *Decoration*). The "Search function" supports wildcards (*), logical operators (and, or, not) and fuzzy search (~).

In this connection, every description has been compiled in accordance with the "grammar" and "syntax" of XML and TEI-P5 electronic text encoding guidelines¹⁶ and contains these key-tags: <msDesc> (manuscript description), which contains a description of a single manuscript; <msIdentifier> (manuscript identifier), which contains the information required to identify the manuscript being described; <msContents> (manuscript contents), which describes the intellectual content of a manuscript or manuscript part; <physDesc> (physical description), which contains a full physical description of a manuscript or manuscript part; <history>, which groups elements describing the full history of a manuscript or manuscript part; <additional>, which groups additional information, combining bibliographic information about a manuscript, or surrogate copies of it with curatorial or administrative information; <msPart> (manuscript part), which contains information about an originally distinct manuscript or part of a manuscript, now forming part of a composite manuscript.

Some of these tags employed in the TEI-P5 scheme for encoding information contained in manuscript descriptions may be associated with tags from the *Dublin Core*¹⁷ standard, which features specific descriptive categories for the description of bibliographical sources.

2. Conclusion

In my opinion, the <e-codices - Virtual Manuscript Library of Switzerland> project represents a useful opportunity for manuscript scholars and librarians to contribute (in a collaborative way) to the growth of research on the fields of Latin Paleography and Codicology: the digitization and codify with XML TEI-P5 standard makes possible the "virtual reunification"³ of medieval and early modern dispersed manuscripts from Swiss libraries, and provides "an immediate Internet access to digital copies of works"³, according to Anne Marie Austenfeld. Today, also for humanities research, digital repositories, digital libraries and virtual libraries are among the most common tools for disseminating digital object collections of cultural, educational, and other kinds of content over the Web: in this context the <e-codices> project, with all its characteristics, is state-of-the-art.

References

1. Bernasconi Reusser M. Libri e manoscritti da biblioteche ticinesi nelle biblioteche digitali e-codices ed e-rara, *Archivio Storico Ticinese* 2012;**151**:156-159.
2. Cusimano F. Le nuove tecnologie per la ricerca umanistica: banche dati, repertori bibliografici, “biblioteche digitali” e “cataloghi aperti dei manoscritti”, *Mediaeval Sophia* 2010;**8**:141-154.
3. Austenfeld AM. Virtual reunification as the future of ‘codices dispersi’: Practices and standards developed by e-codices – Virtual Manuscript Library of Switzerland, *IFLA Journal* 2010;**36**(2):145–154.
4. Tammaro AM. Interoperabilità. Verso un modello aperto, *Biblioteche oggi* 2007;**6**:55-58.
5. Zillur R. Digital Libraries in the New Millennium, *Bangladesh Journal of Library and Information Science* 2012;**2**(1)103-111.
6. Ridi R. La biblioteca digitale: definizioni, ingredienti e problematiche, *Bollettino AIB* 2004;**44**(3):273-344.
7. Karadkar UP et al, Metadocuments supporting digital library information discovery, *International Journal on Digital Libraries* 2004;**4**:25-30.
8. Gonçalves MA, Fox E, Watson LT. Towards a digital library theory: a formal digital library ontology, *International Journal on Digital Libraries* 2008;**8**:91-114.
9. Barbero G. Herzog August Bibliothek: manoscritti e XML, *Biblioteche oggi* 2006;**1**:pp. 64-66
10. Pierazzo E. *La codifica dei testi: un'introduzione*, Roma: Carocci, 2005.
11. Ciotti F, editor. *Burnard L., Sperberg-McQueen C. M., Il manuale TEI Lite. Introduzione alla codifica elettronica dei testi letterari*, Milano: Sylvestre Bonnard, 2005.
12. Rusty Harold E, Scott Means W. *XML. Guida di riferimento*, Milano: Apogeo, 2001, pp. 3-4.
13. *TEI P5 Guidelines for Electronic Text Encoding and Interchange by the TEI Consortium. Originally edited by Sperberg-McQueen CM and Burnard L for the ACH-ALLC-ACL Text Encoding Initiative. Now entirely revised and expanded under the supervision of the Technical Council of the TEI Consortium. Burnard L, Bauman S, editors. 2.3.0. Last updated on 17th January 2013*, Charlottesville, Virginia, Text Encoding Initiative Consortium, 2013, pp. XXXIII-LII.
14. Elena T et al. Historical research in archives: user methodology and supporting tools, *International Journal on Digital Libraries* 2010;**11**:25-36.
15. Burkard B, Vogeler G, Gruner S. Informatics for Historians: Tools for Medieval Document XML Markup, and their Impact on the History-Sciences, *Journal of Universal Computer Science* 2007;**14**(2):193-210.
16. Fradejas Rueda JM. La codificación XML/TEI de textos medievales, *Memorabilia* 2009-2010;**12**:219-247.
17. Gorman M. Metadati o catalogazione? Una falsa alternativa, *Biblioteche oggi* 2001;**5**:14.